DOCUMENT RESUME

ED 341 710 TM 017 889

AUTHOR Forsyth, G. Alfred; And Others
TITLE Factors Affecting the Misinterpretation of Research.
PUB DATE Aug 91
NOTE 38p.; Paper presented at the Annual Meeting of the American Psychological Association (99th, San Francisco, CA, August 16-20, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Data Interpretation; Economics; Foreign Countries; Generalization; Higher Education; Knowledge Level; Mathematics; Methods Courses; *Misconceptions; Multiple Regression Analysis; Psychology; Research Methodology; Research Problems; *Research Reports; Research Utilization; Sampling; *Scientific Concepts; Student Attitudes; *Undergraduate Students
IDENTIFIERS Bloomsburg University PA; *Cause Effect Relationships; Dependent Variables; Millersville University PA; *Misinterpretation; University of Victoria BC

ABSTRACT

To determine the factors that influence interpretation of a research study description, 8 descriptions of research studies were interpreted by 320 undergraduate students in psychology, mathematics, and economics statistics and research methods classes at the University of Victoria (British Columbia, Canada), Millersville University (Pennsylvania), or Bloomsburg University (Pennsylvania). Results of stepwise multiple regression analyses indicate that students in all three disciplines base judgments of the generalizability of research findings and confidence in drawing cause-and-effect conclusions on their belief that there is an independent-dependent variable relationship. Students attended to neither the existence of random sampling in generalizing findings, nor the use of random assignment in drawing cause-and-effect conclusions. Students in second-level psychology courses and lower-level mathematical statistics courses incorrectly altered their confidence in drawing cause-and-effect conclusions based on the statistic used rather than the research methods reported. A research methods taxonomy is presented for overcoming these errors in interpreting research. Greater emphasis on interpretation-of-research assignments in statistics and methods courses is recommended. Three tables, 16 figures, and a 10-item list of references are included. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
G. ALFRED FORSYTH
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED341710

FACTORS AFFECTING THE MISINTERPRETATION OF RESEARCH

G. Alfred Forsyth
(Millersville University)

Peter H. Bohling
(Bloomsburg University)

Richard B. May
(University of Victoria)

Poster Session Paper at the 99th Annual Meeting of the American Psychological Association, San Francisco, California, August 1991

TM017886

BEST COPY AVAILABLE

ABSTRACT

THE GOAL OF THIS STUDY WAS TO DETERMINE WHAT FACTORS INFLUENCE THE INTERPRETATION OF A RESEARCH STUDY DESCRIPTION.

EIGHT DESCRIPTIONS OF RESEARCH STUDIES WERE INTERPRETED BY STUDENTS IN PSYCHOLOGY, MATHEMATICS, AND ECONOMICS STATISTICS AND RESEARCH METHODS CLASSES AT THREE UNIVERSITIES. RESULTS OF STEPWISE MULTIPLE REGRESSION ANALYSES OF THE DATA INDICATE THAT STUDENTS IN ALL THREE DISCIPLINES BASE JUDGMENTS OF GENERALIZABILITY OF FINDINGS AND CONFIDENCE IN DRAWING CAUSE AND EFFECT CONCLUSIONS ON THEIR BELIEF THAT THERE IS AN INDEPENDENT-DEPENDENT VARIABLE RELATIONSHIP.

STUDENTS DID NOT ATTEND TO THE EXISTENCE OF RANDOM SAMPLING IN GENERALIZING FINDINGS NOR THE USE OF RANDOM ASSIGNMENT IN DRAWING CAUSE AND EFFECT CONCLUSIONS.

STUDENTS IN SECOND-LEVEL PSYCHOLOGY COURSES INCORRECTLY ALTERED THEIR CONFIDENCE IN DRAWING CAUSE AND EFFECT CONCLUSIONS BASED ON THE STATISTIC USED TO ANALYZE THE DATA RATHER THAN RESEARCH METHODS REPORTED. A SIMILAR MISUNDERSTANDING ABOUT THE RELATIONSHIP OF THE CORRELATION COEFFICIENT AND DRAWING CAUSE AND EFFECT CONCLUSIONS WAS EVIDENT FOR THE STUDENTS IN THE LOWER-LEVEL MATHEMATICAL STATISTICS COURSES.

A RESEARCH METHODS TAXONOMY FOR OVERCOMING THESE INTERPRETATION-OF-RESEARCH ERRORS IS PRESENTED. GREATER EMPHASIS ON INTERPRETATION-OF-RESEARCH ASSIGNMENTS IN STATISTICS AND RESEARCH METHODS COURSES IS RECOMMENDED.

Factors Affecting the Misinterpretation of Research

During the past few years, concern has been expressed about the inability of consumers of literature to interpret correctly psychological research articles. For example, May and Hunter (1988) reported that less than one-third of students and faculty correctly interpret research studies in the areas of statistical significance, cause and effect, and generalizability of results. Rosnow and Rosenthal (1989) emphasize that a major problem associated with the interpretation of research is too great a reliance on statistical significance. Bohling and Forsyth (1990) describe three additional problems in the interpretation of behavioral science research: 1) inattention to random sampling and random assignment research methods; 2) an incorrect understanding of the relationship between the correlation statistic and cause-and-effect conclusions; and 3) an overreliance on the interpreter's beliefs about specific independent-dependent variable relationships.

As noted by May and Hunter (1990), statistical testing is normally introduced using the normal curve or random sampling model in statistics and research methods books in psychology. A review of statistics and research methods books in mathematics and in business and economics indicates that this overreliance on the random sampling model is present in these disciplines as well as psychology.

Many of the textbooks include sections on the importance of random sampling in research and make it explicit that the statistical procedures covered in the book require random samples. As Glenberg (1988) points out, random samples are rarely obtained in real-world investigations. Without random sampling from some specified population of interest, inferential statistical procedures are inappropriate and generalization through inference is inappropriate. The alternative of using permutation tests on scores as well as ranks is present in some newer statistics and

research methods books (e.g., May, Masson and Hunter, 1990, and Forsyth and Bohling, 1991). However, most books still limit the permutation model to rank-order data and do not present it as an alternative with interval or ratio data in studies in which a random sampling method is not used. One goal of the present study is to examine how the presence or absence of random sampling affects students' confidence in generalizing the results of a study.

Just as random sampling is an important research method to permit inferences or generalizations to the population of interest, random assignment is an important research method for drawing cause and effect conclusions about the independent-dependent variable relationship. Most statistics and research methods students learn that random assignment to treatment conditions is important if one is to conclude that differences in the dependent variable are caused in part by the independent variable. However, as Stevens (1990) points out, most students in these courses do not make connections between course material and reading the literature because courses focus on mathematics and techniques with too little training in critically evaluating literature. If students do make such connections, they should be more confident in drawing cause and effect conclusions when research has involved random assignment rather than using a non-experimental design. A second purpose of the present study is to determine the degree to which students across disciplines attend to random assignment vs. classificatory research methods in determining the appropriateness of drawing a cause and effect conclusion when interpreting a study.

An examination of recently published statistics and research methods books in psychology and mathematics indicates that most include a section that warns students of the inappropriateness of drawing a cause and effect conclusion if the study is correlational. Unfortunately, most of the books do not make a clear distinction between a correlational study being one in which there is not a manipulated independent variable vs. it being a study in which a correlation coefficient is computed in analyzing the data. Only a few books (e.g., Keppel and Zedeck, 1989) clearly articulate that it is not the use of the correlation coefficient statistic that

lowers confidence in drawing a cause and effect conclusion. Instead, they point out that a classificatory or correlational research method should reduce confidence in drawing cause and effect conclusions. A third purpose of the present research is to determine if the use of r vs. t to report results affect drawing cause and effect conclusions. Based on our review of available books, we hypothesize that students in psychology and mathematics statistics and research methods courses will erroneously lower their confidence in drawing a cause and effect conclusion when a correlation coefficient is used to determine statistical significance instead of a t.

Instead of attending to the research methods in interpreting a research study, we believe that the availability-heuristic reasoning process is used in interpreting research studies. That is, individuals will believe that there is an independent-dependent variable relationship if they have first-hand knowledge of referent examples that involve those variables. If the results of a study are in accord with the belief resulting from the availability-heuristic reasoning process, individuals will be confident that the results of the study are due to a cause and effect relationship and can be generalized. If the results of a study are not consistent with availability-heuristic-based beliefs, individuals will reject a cause and effect interpretation of the study and rate generalization of results as low. For example, in reading an experimental study that indicates a significant relationship between watching violence and emitting physically aggressive behavior, an individual may rate generalization and cause and effect low because she knows a person who never watched violence on T.V., yet killed three people. Relatedly, she might indicate that she knows another person who watches a great deal of T.V. violence and never exhibits aggressive behavior. Thus, another goal of the current study is to determine the degree to which individuals' beliefs about the existence of an independent-dependent variable relationship correlates with confidence in generalizing the results of a study or in drawing cause and effect conclusions.

The research interpretation survey is designed to examine the relative importance of subjects' belief, random sampling, and random assignment in the

interpretation of findings being generalizable and supporting cause and effect conclusions. It also permits an assessment of the degree to which the type of statistical analysis affects confidence in drawing cause and effect conclusions.

Method

A total of 320 undergraduate students in statistics and research methodology courses in either the departments of psychology, mathematics, or economics at either Millersville University, Bloomsburg University, or University of Victoria read eight brief descriptions of research studies and answered questions about each study. The students are from one of two levels of statistics and research methodology courses in their discipline. Two of the universities are state-owned universities in Pennsylvania and the third is a province-owned university in British Columbia.

The eight studies read by each student consist of two studies from each of four cells created by the two-by-two crossing of random sampling vs. available subjects and random assignment vs. existing classificatory groups. For example, a study involving available groups of children who had been read to either very little or a great deal is in the nonrandom sampling and nonrandom assignment cell of the two-by-two matrix defining the four types of studies. An example of a study involving both random sampling and random assignment is one in which a random sample of obese individuals (50 pounds overweight) from a large health care organization are randomly assigned to one of two different treatment plans with one-year weight change as the dependent variable. One of the studies involving a random sample of subjects from existing classificatory groups is a psychometric study in which a new test designed to predict college success is given to all high school applicants applying to a university. One random sample of students with test scores between the 30th and 40th percentile ranks is randomly selected for admission to the university. A second sample of students with test scores between the 70th and 80th percentile ranks is randomly selected for admission to the university. The cumulative grade point averages of the two groups is compared after the first academic year. Finally, an example of a study

involving available subjects who are randomly assigned to treatment conditions is one in which volunteers from a local civic organization are randomly assigned to consume either alcoholic or non-alcoholic drinks prior to reaction-time testing in a simulated driving apparatus. The description of each study concludes with a statement about which group was higher on the dependent variable and "$p < .02$" is in parenthesis directly after that statement.

After reading each brief study description, subjects rate their personal belief in there being a meaningful relationship between the independent and dependent variables. Also, using a six-point scale, they rate their confidence that: 1) the results are statistically significant; 2) the results of the study are generalizable; 3) a cause and effect conclusion is appropriate; 4) a cause and effect conclusion would be appropriate if the data were analyzed with a correlation coefficient; and 5) a cause and effect conclusion would be appropriate if the data were analyzed with a *t*-test.

The questions are written within the context of the research study. For example, in the alcohol and simulated driving reaction-time study, the rating question for generalizability is "Based on the research methodology, how confident are you in generalizing the result of this study to other drivers?" The rating responses are from "1" representing not at all confident to "6" representing extremely confident. Similarly, the confidence in drawing a cause and effect conclusion is phrased within the context of the study. Again using the alcohol and reaction-time study as an example, subjects responded to the following rating question: Based on the research methodology, how confident are you that a cause and effect conclusion is appropriate about the alcohol consumption and reaction time variables?

The purpose of beginning the generalizability and cause and effect questions with "Based on the research methodology . . ." is to direct the subjects' attention to the research method in the description. The survey booklets present the survey on the left side and the rating questions on the right side of each open page. This arrangement allows the subject to go back to the study description to examine the

research method while answering the rating questions. All of this is intended to encourage students to interpret studies based on the method used.

Subjects were invited to follow-up sessions to learn more about the purpose of the research and our research findings. Teachers of the courses were provided with a brief summary of the findings. The teachers provided course syllabi. These syllabi along with the catalogue course descriptions permit an assessment of the subjects' familiarity with statistical techniques (e.g., correlation, *t*-tests, and analysis of variance) and the inclusion of random sampling and random assignment topics in the course.

Results

The data for each class are analyzed separately. A stepwise multiple regression analysis for each class examines the relative importance of the random sampling, random assignment, and belief variables in accounting for variability in the confidence in generalizability ratings. If students base their generalizability ratings on inferences appropriate to the research method, random sampling should be the first variable accounting for the greatest variability in the stepwise multiple regression analysis.

The first stepwise multiple regression analysis addresses the question of what variables account for variation in generalizability ratings. Table 1 presents the percents of generalizability variation that is accountable by the random sampling, random assignment and belief variables. As shown in Table 1, the belief variable accounts for more variation than random sampling or random assignment in 13 of the 14 classes. The belief variable is significant at $p < .0001$ in all 14 classes. Going beyond significance, the belief variable accounts for a large percent of the generalizability variation in all but one class. In contrast, neither random sampling nor random assignment account for more than 5 percent of the generalization variability for the 13 classes in which belief vs. the first variable.

The large r^2 values for the belief predictor would have been expected to be in the random sampling column if students were more sensitive to the importance of random sampling for generalizing the results of a study. These ideal results are presented in Figure 1 using r rather than r^2 to reflect the strength of relationship. Figures 2, 3, 4, and 5 provide a contrast of how different the obtained pattern of results is from the Figure 1 ideal pattern. Figures 2, 3, 4, and 5 show the correlation coefficients of random sampling, random assignment, and belief with generalizability ratings for four of the 14 classes.

In the one class in which belief is not the first variable in the stepwise multiple regression analysis, random sampling accounts for 16 percent of the generalizability variation and belief accounts for an additional 10 percent. In that class, the mean generalizability rating for the four research descriptions involving random sampling is 3.53 compared to a 2.33 mean generalizability for the four nonrandom-sample research descriptions. The overall average generalizability rating for the four random-sampling research descriptions is 3.88 for the other 13 classes. This compares to a 3.57 overall average generalizability rating for the four nonrandom-sampling research descriptions for these same 13 classes.

Table 1. Percents of the Generalizability Variation Accountable by the Random Sampling, Random Assignment and Belief Variables in the Stepwise Multiple Regression Analysis.

	Predictor		
Course/University	Random Sampling	Random Assignment	Belief
Psychology 1, U1	01	02*	44**
Psychology 2, U1	00	01	50**
Mathematics 1, U1	00	01	34**
Mathematics 2, U1	00	00	37**
Economics 1, U1	00	00	28**
Economics 2, U1	02	02	58**
Psychology 1, U2	00	00	58**
Psychology 2, U2	03*	01	47**
Mathematics 1, U2	01	00	45**
Economics 1, U2	00	00	30**
Economics 2, U2	05*	02	47**
Psychology 2, U3	16**	00	10**
Mathematics 2, U3	00	00	37**
Economics 2, U3	01	03*	55**

* $p < .01$

** $p < .0001$

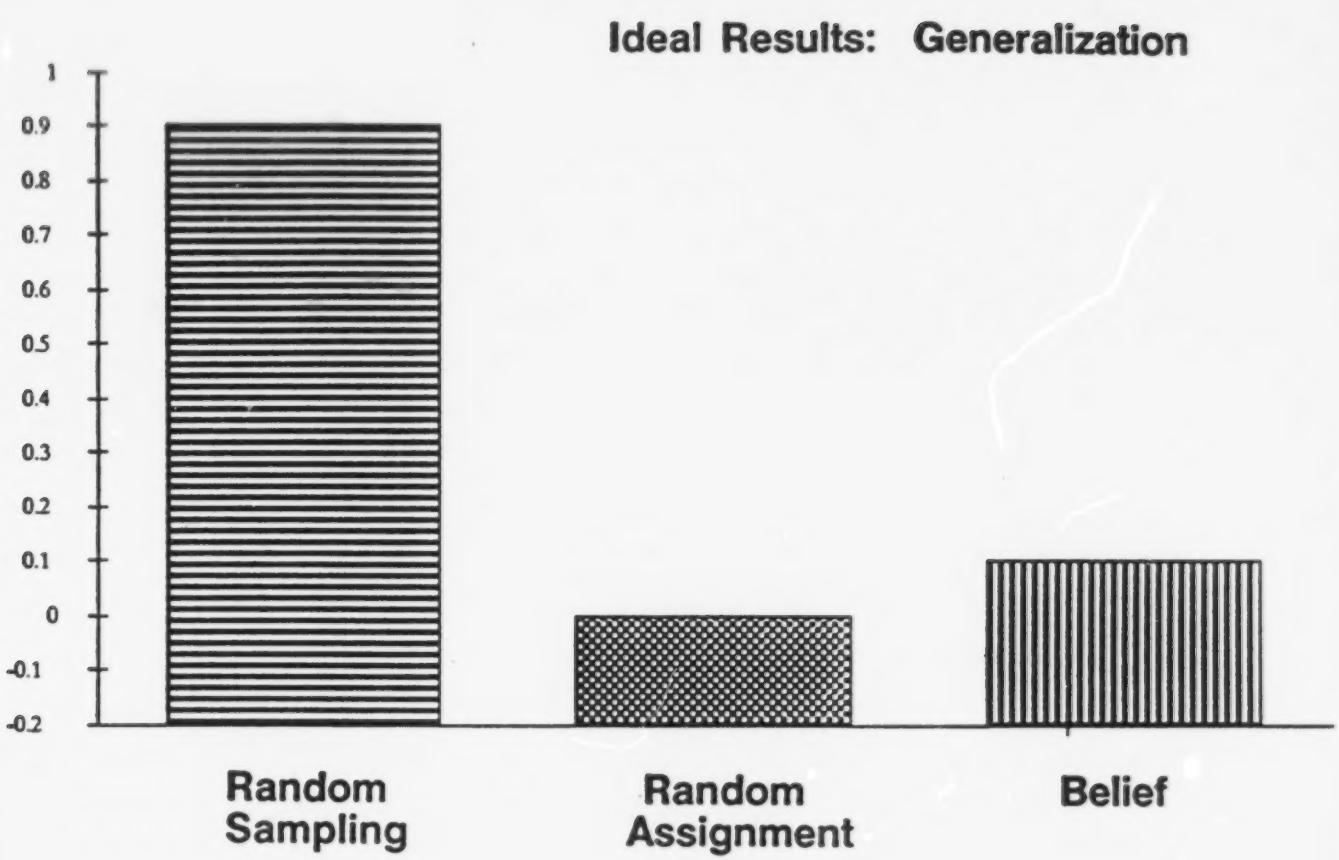

Figure 1. **Ideal Correlations of Confidence in Generalizations with Random Sampling, Random Assignment, and Belief Independent Variables**

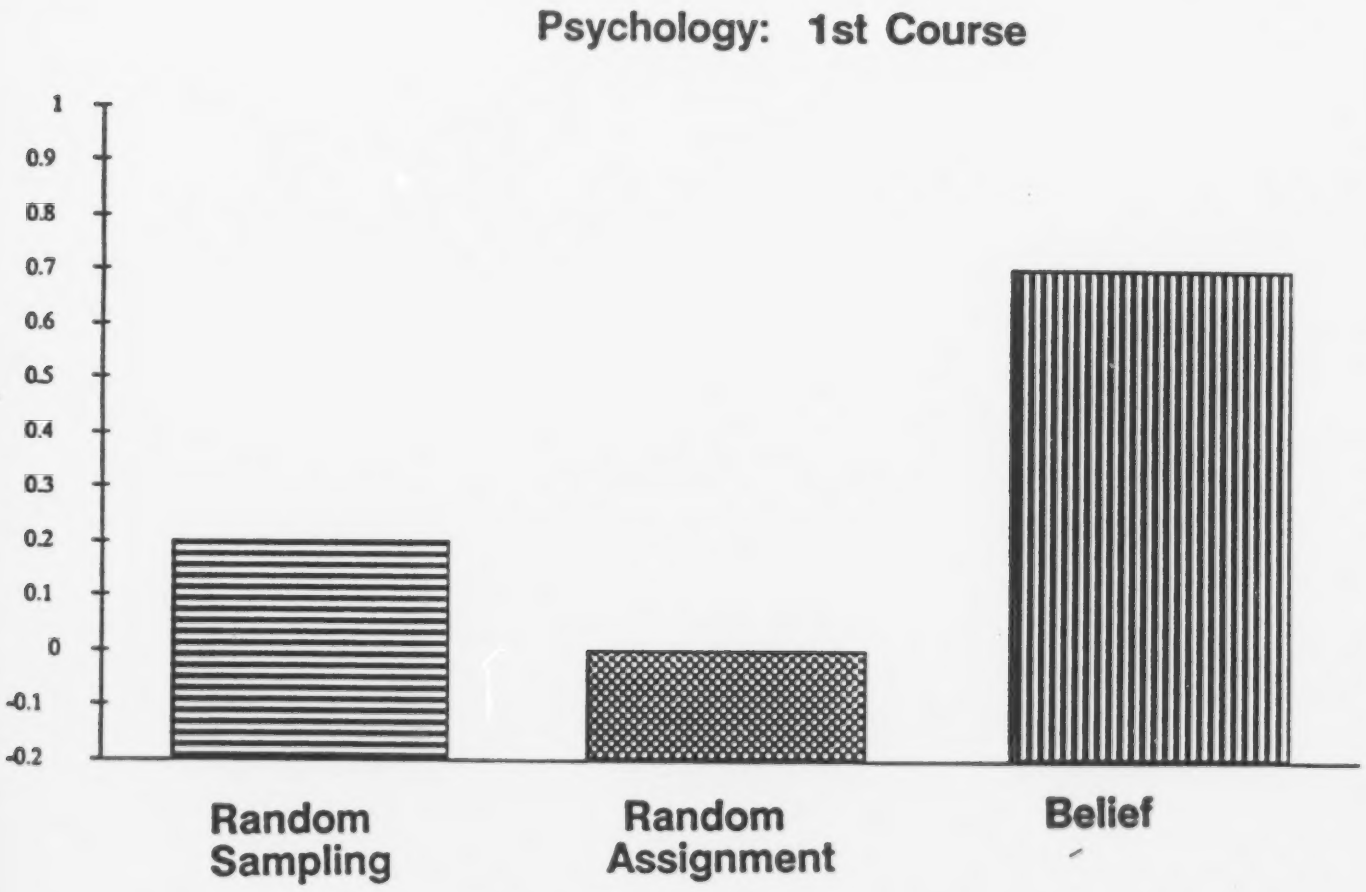

Figure 2. **Correlations of Confidence in Generalizations with Random Sampling, Random Assignment, and Belief Variables: Psychology Level 1**

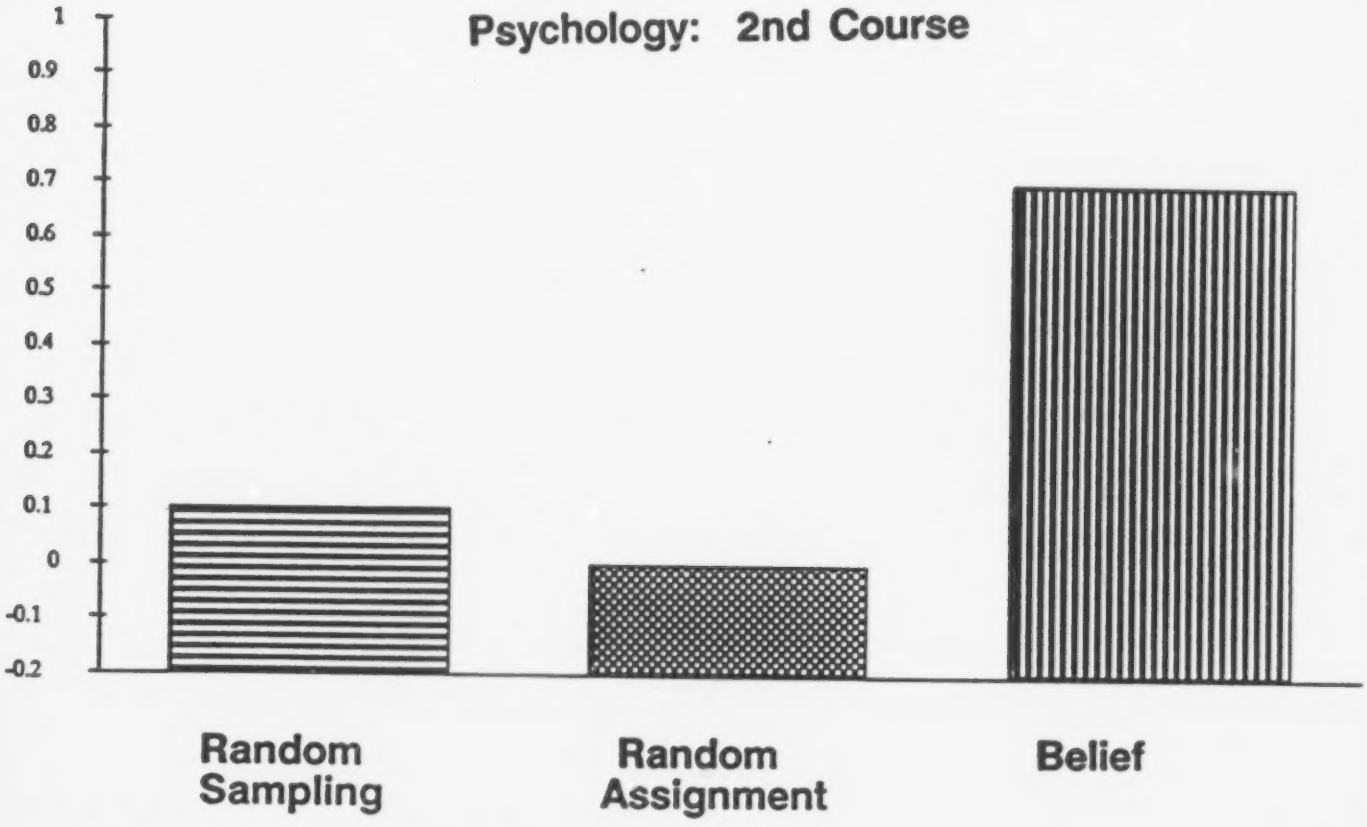

Figure 3. **Correlations of Confidence in Generalizations with Random Sampling, Random Assignment, and Belief Variables: Psychology Level 2**

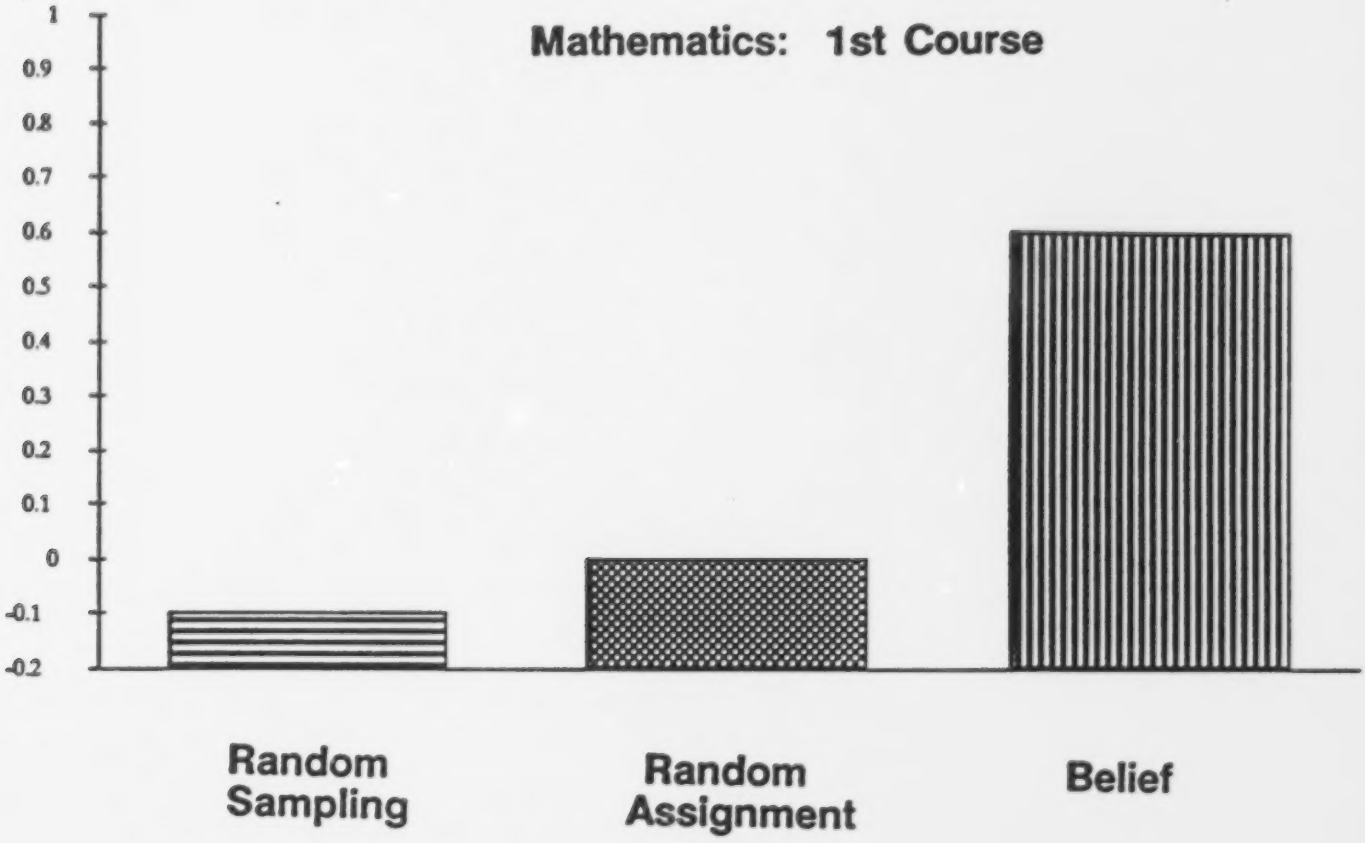

Figure 4. **Correlations of Confidence in Generalizations with Random Sampling, Random Assignment, and Belief Variables: Mathematics Level 1**

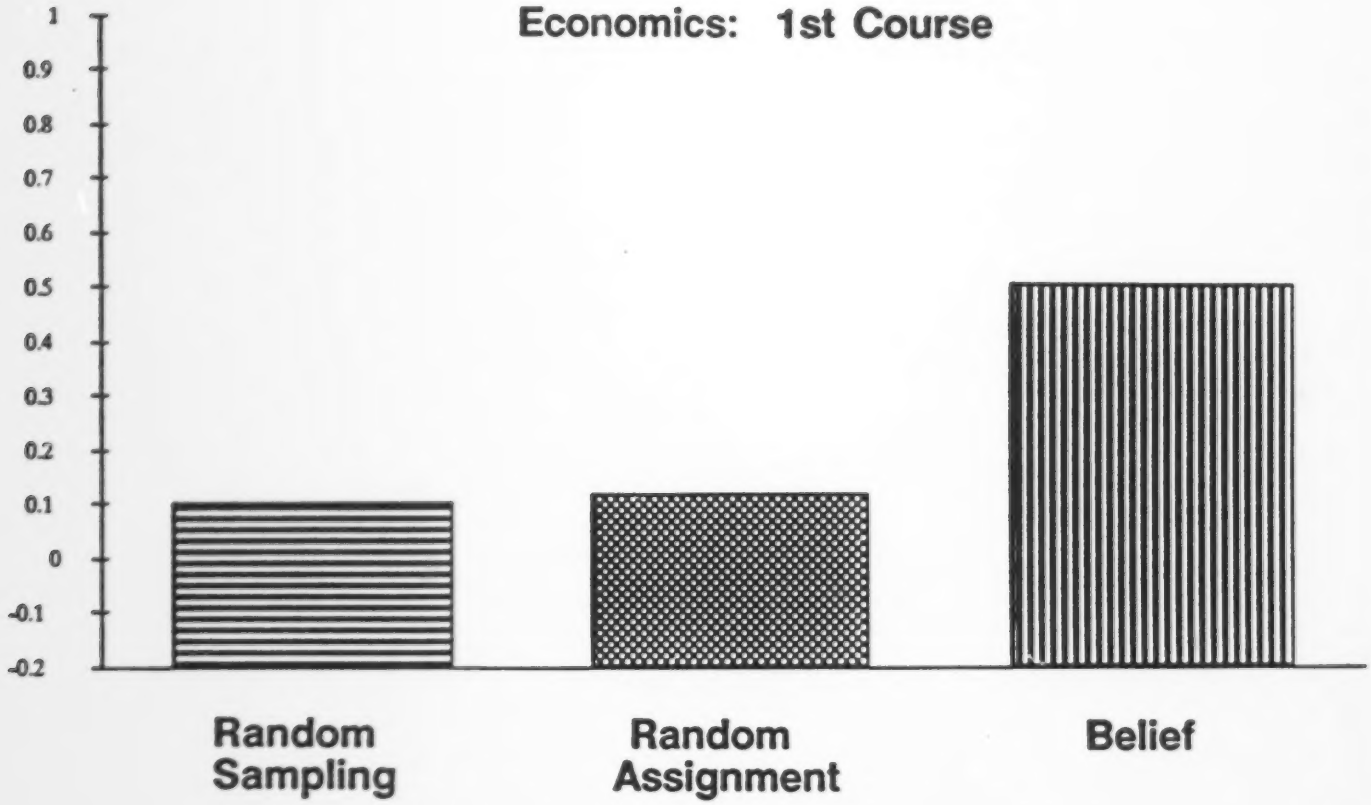

Figure 5. **Correlations of Confidence in Generalizations with Random Sampling, Random Assignment, and Belief Variables: Economics Level 1**

The second stepwise multiple regression analysis addresses the question of what variables account for variation in cause and effect ratings. Table 2 presents the percents of cause and effect variation that is accountable by the random sampling, random assignment, and belief variables. As shown in Table 2, the belief variable again accounts for more variation than random sampling or random assignment in 13 of the 14 classes. The belief variable is significant at $p < .0001$ in all 14 classes. Going beyond significance, Table 2 shows that belief accounts for 49 percent ($r = .7$) or more of the variation in cause and effect ratings for nine of the 14 classes. In one class, random assignment is the first variable and accounts for 9 percent of the cause and effect variation. Neither random sampling or random assignment accounts for more than 2 percent of the cause and effect variation in the other 13 classes.

In the case of confidence in cause and effect ratings, the large r^2 values for the belief predictor would have been expected to be in the random assignment column if students understood the importance of random assignment research method for drawing cause and effect conclusions. These ideal results are presented in Figure 6 using r rather than r^2 to reflect the strength of relationship. Figures 7, 8, 9, and 10 illustrate how different the obtained patterns of results are from the Figure 6 ideal pattern. Figures 7, 8, 9, and 10 show the correlation coefficients of random sampling, random assignment, and belief with cause and effect ratings for four of the 14 classes.

The one class in which random assignment accounts for more of the cause and effect variation than belief is the same class in which random sampling accounts for more generalizability variation than belief. For this one psychology class, the stepwise multiple regression analysis shows that random assignment accounts for 9 percent of the cause and effect variation and that belief contributes an additional 6 percent. In that class, the mean cause and effect rating for the four research descriptions involving random assignment is 3.47 compared to a 2.61 mean cause and effect rating for the four nonrandom-assignment research descriptions. The overall average cause and effect rating for the four random-assignment research descriptions is 4.13 for the

other 13 classes. This compares to a 3.80 overall average cause and effect rating for the four nonrandom-assignment research descriptions for these 13 classes.

Table 2. Percents of the Cause and Effect Variation Accountable by the Random Sampling, Random Assignment, and Belief Variables in the Stepwise Multiple Regression Analysis.

	Predictor		
Course/University	Random Sampling	Random Assignment	Belief
Psychology 1, U1	00	00	60**
Psychology 2, U1	00	00	52**
Mathematics 1, U1	00	00	44**
Mathematics 2, U1	01	00	50**
Economics 1, U1	00	00	51**
Economics 2, U1	00	00	64**
Psychology 1, U2	00	01	63**
Psychology 2, U2	01	00	30**
Mathematics 1, U2	00	02*	54**
Economics 1, U2	00	00	51**
Economics 2, U2	00	00	34**
Psychology 2, U3	02	09**	06**
Mathematics 2, U3	00	00	49**
Economics 2, U3	00	00	40**

* $p < .01$

** $p < .0001$

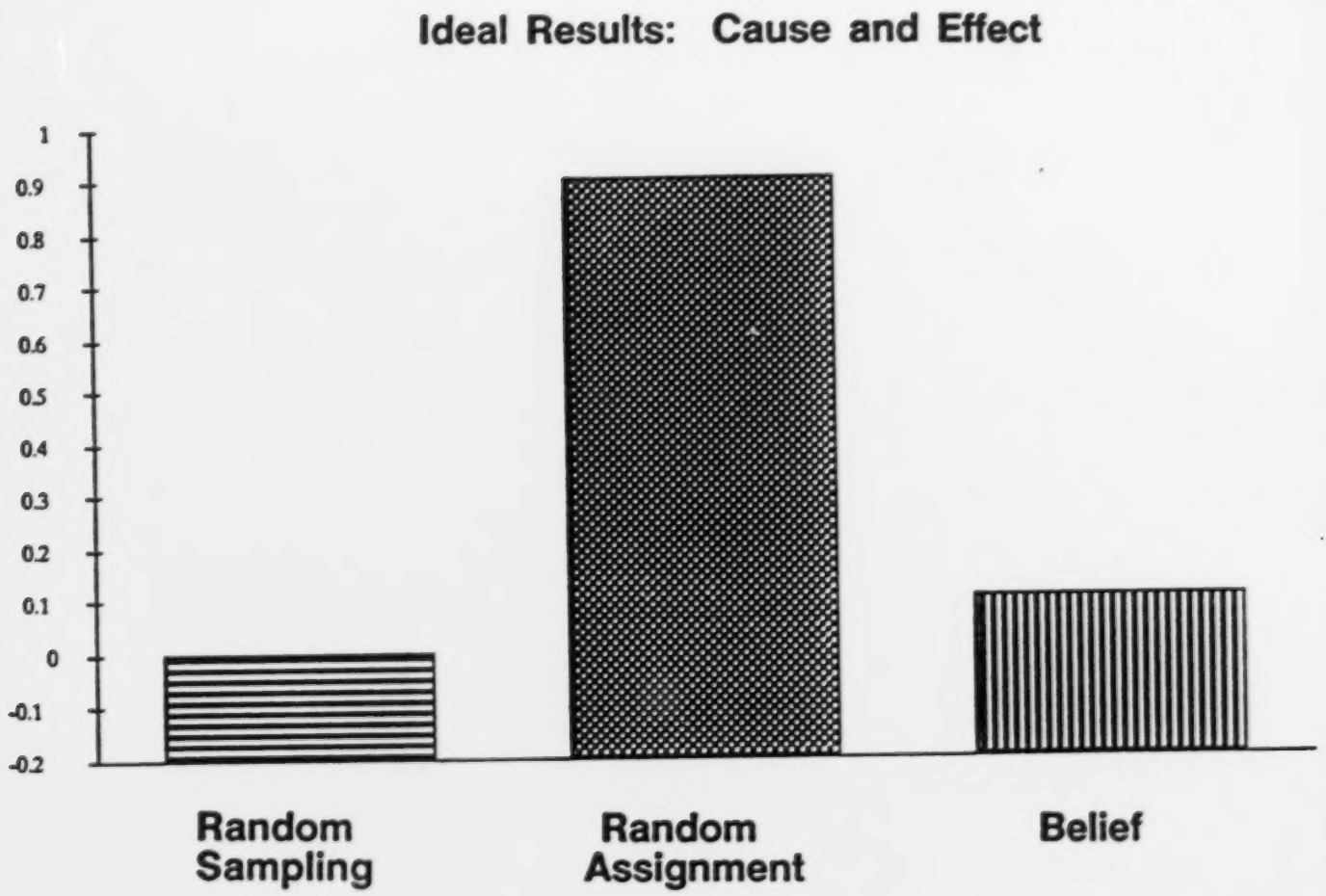

Figure 6. **Ideal Correlations of Confidence in Cause and Effect Conclusions with Random Sampling, Random Assignment, and Belief Independent Variables**

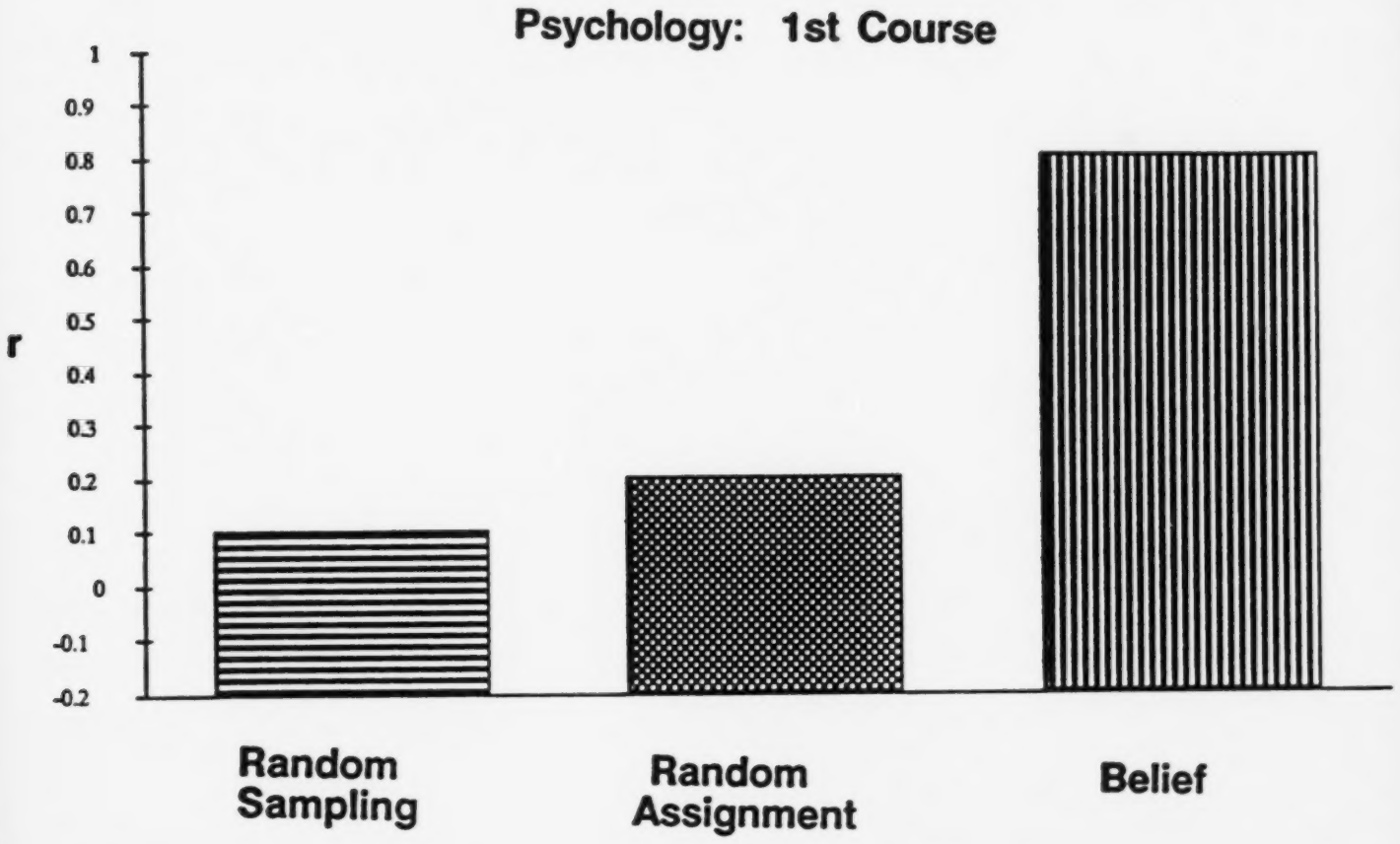

Figure 7. **Correlations of Confidence in Cause and Effect Conclusions with Random Sampling, Random Assignment, and Belief Variables: Psychology Level 1**

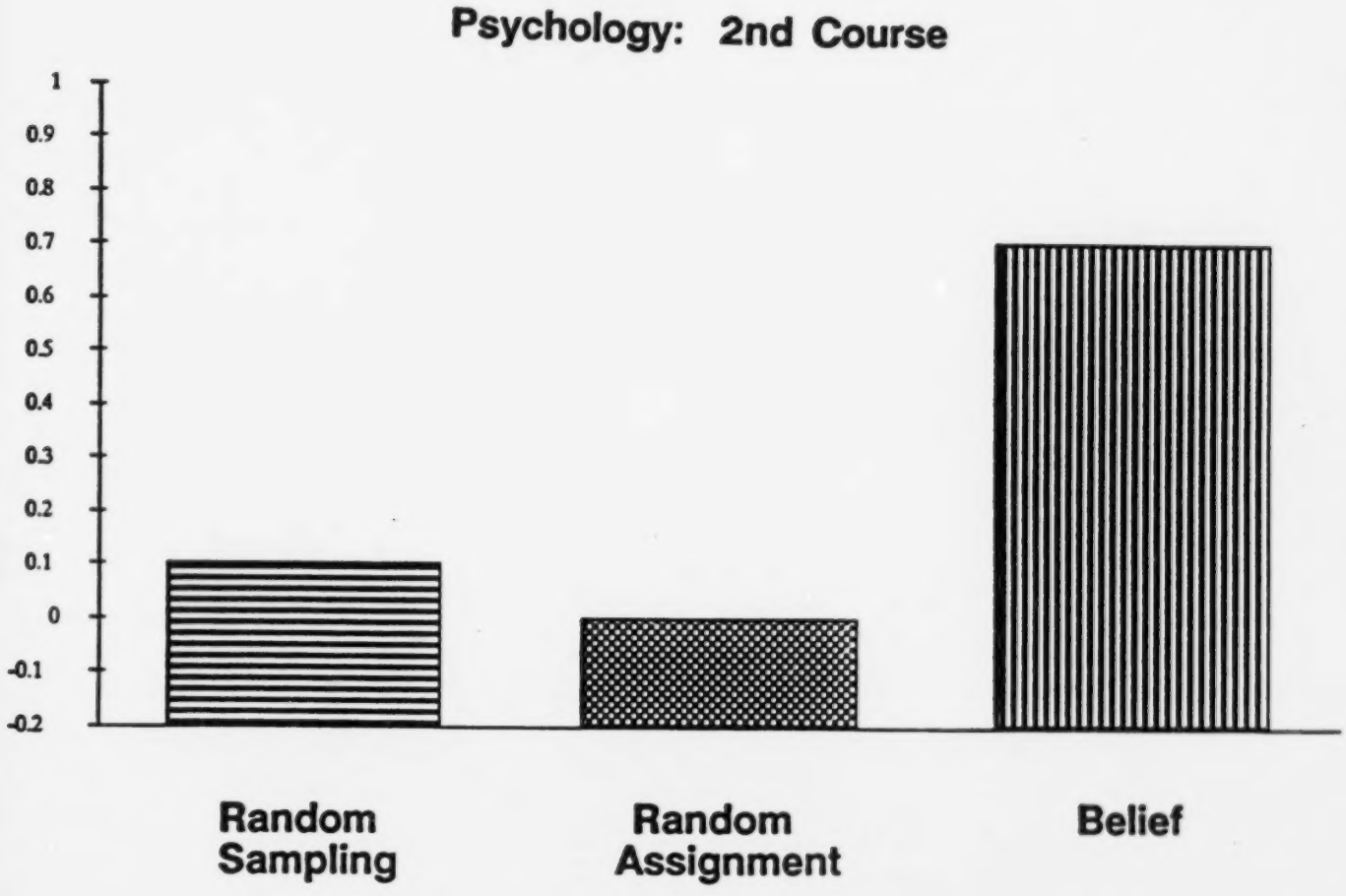

Figure 8. **Correlations of Confidence in Cause and Effect Conclusions with Random Sampling, Random Assignment, and Belief Variables: Psychology Level 2**

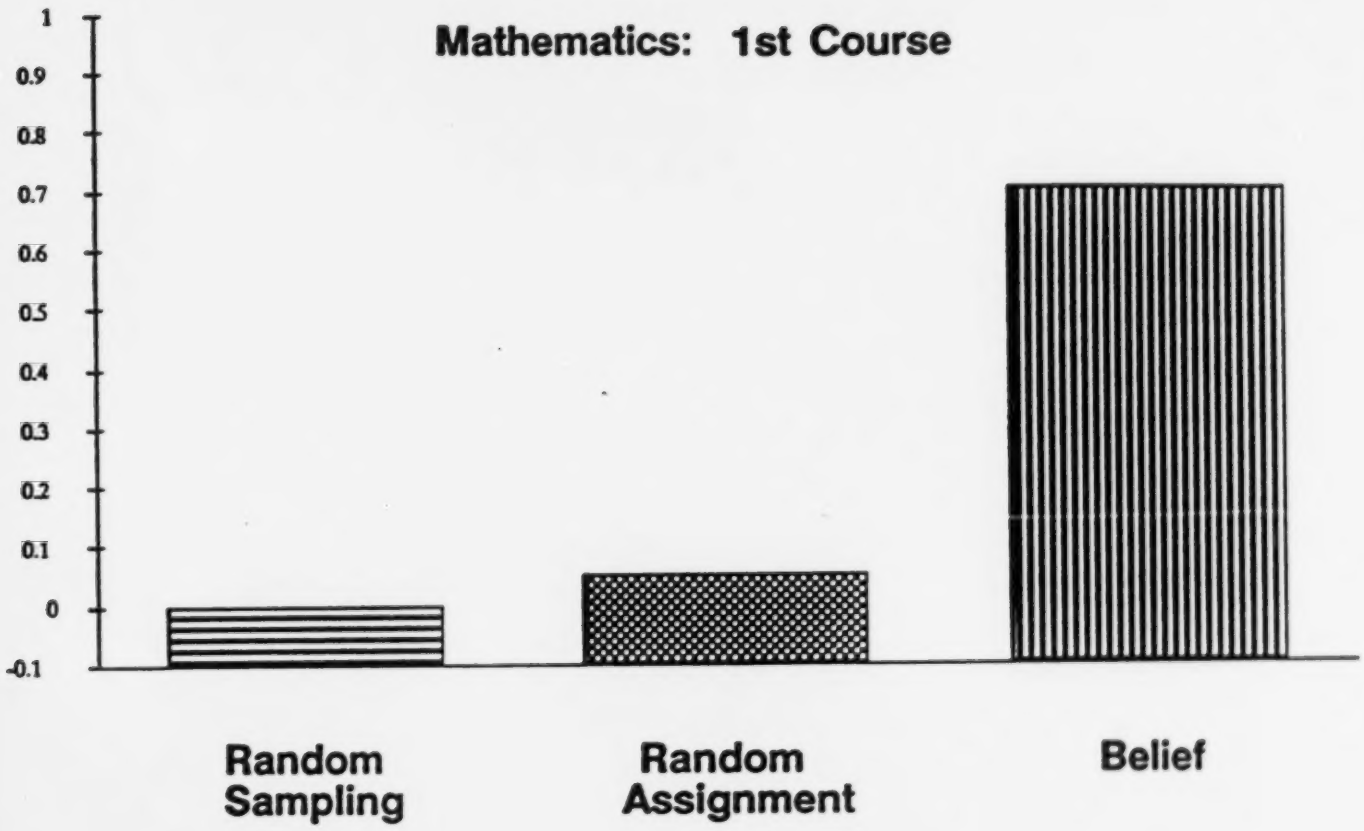

Figure 9. **Correlations of Confidence in Cause and Effect Conclusions with Random Sampling, Random Assignment, and Belief Variables: Mathematics Level 1**

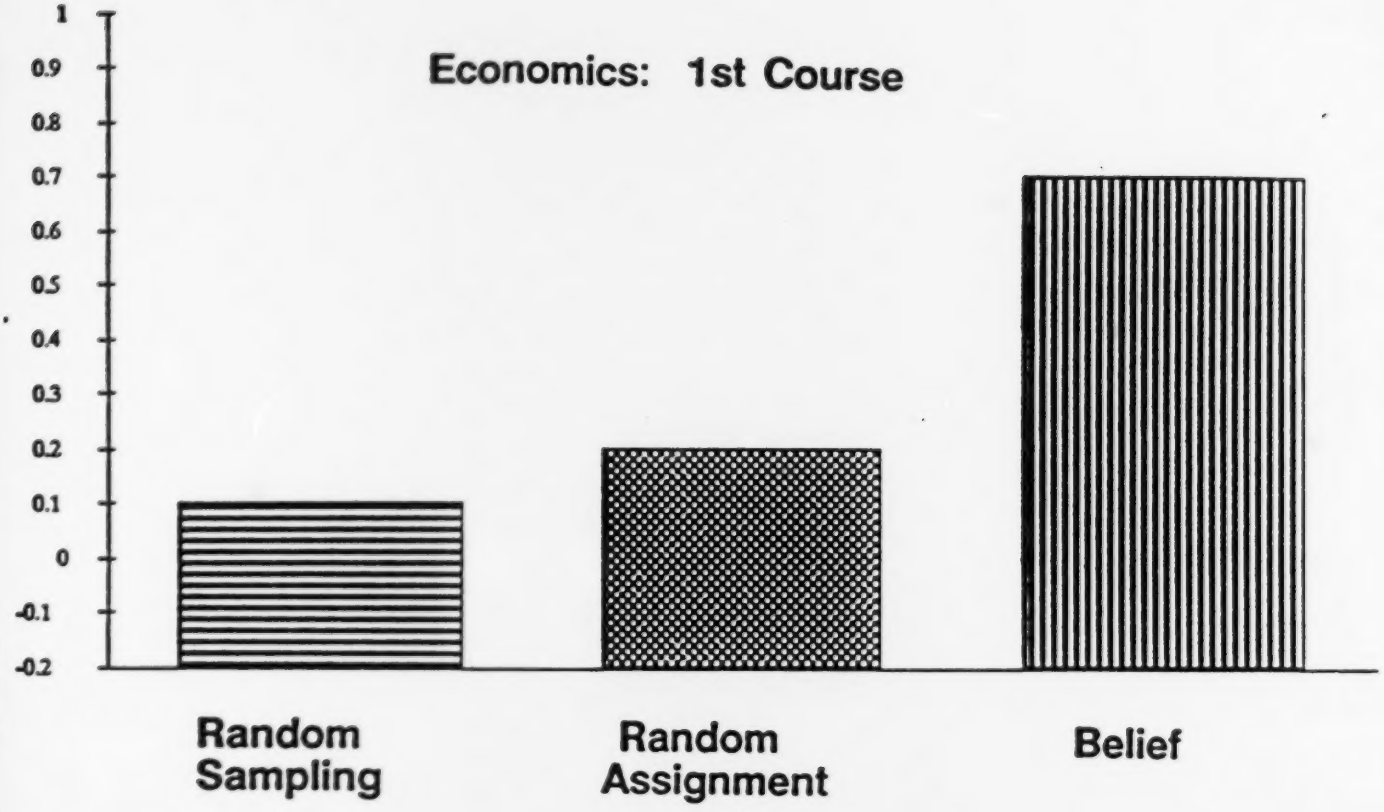

Figure 10. Correlations of Confidence in Cause and Effect Conclusions with Random Sampling, Random Assignment, and Belief Variables: Economics Level 1

The third set of analyses examines whether or not confidence in cause and effect is affected by the type of statistical analysis performed. Table 3 presents the mean confidence in cause and effect ratings for random assignment studies when significance is reported with a correlation coefficient vs. with a *t*-test. The significance column in Table 3 indicates for which classes a repeated-measures analysis shows a significant difference between cause and effect confidence when a correlation coefficient is used to analyze data and cause and effect confidence when a *t*-test is used to analyze the data.

Students in the following classes rate their confidence in a cause and effect conclusion significantly higher ($p < .01$) when study results are reported with a *t*-test than when study results are reported with a correlation coefficient: psychology level 2 at university 1, psychology level 2 at university 2, psychology level 2 at university 3, mathematics level 1 at university 1, mathematics level 2 at university 1, mathematics level 1 at university 2, and mathematics level 2 at university 3. This greater confidence in cause and effect with a *t*-test than with a correlation is not evident in the two first-level psychology statistics and research methods classes. Indeed, for psychology level 1 at university 1, cause and effect confidence was significantly higher with a correlation rather than a *t*-test. Confidence in cause and effect is not significantly affected by the type of statistical analysis performed for any of the five economics statistics and research methods classes.

Because the four research descriptions used in the Table 3 analyses all involved random assignment, ideally there should not have been a difference based on data being analyzed with a correlation statistic vs. a *t*-test. These ideal results are presented in Figure 11. Figures 12, 13, 14, and 15 are graphic presentations of the cause and effect ratings with correlation vs. a *t*-test for four of the classes.

Table 3. Mean Confidence in Cause and Effect Ratings for Random Assignment Studies when Significance is Reported with a Correlation Coefficient vs. with a t-test.

Course/University	Statistical Test: Correlation Coefficient	Statistical Test: t-test	Significance
Psychology 1, U1	3.74	3.49	*
Psychology 2, U1	2.79	4.31	**
Mathematics 1, U1	3.38	4.23	**
Mathematics 2, U1	3.68	4.07	*
Economics 1, U1	4.15	4.17	–
Economics 2, U1	4.06	3.94	–
Psychology 1, U2	3.88	3.98	–
Psychology 2, U2	3.15	4.10	**
Mathematics 1, U2	3.56	4.08	*
Economics 1, U2	4.39	4.23	–
Economics 2, U2	4.00	4.05	–
Psychology 2, U3	2.34	3.21	**
Mathematics 2, U3	3.65	4.37	**
Economics 2, U3	4.05	4.30	–

* $p < .01$

** $p < .0001$

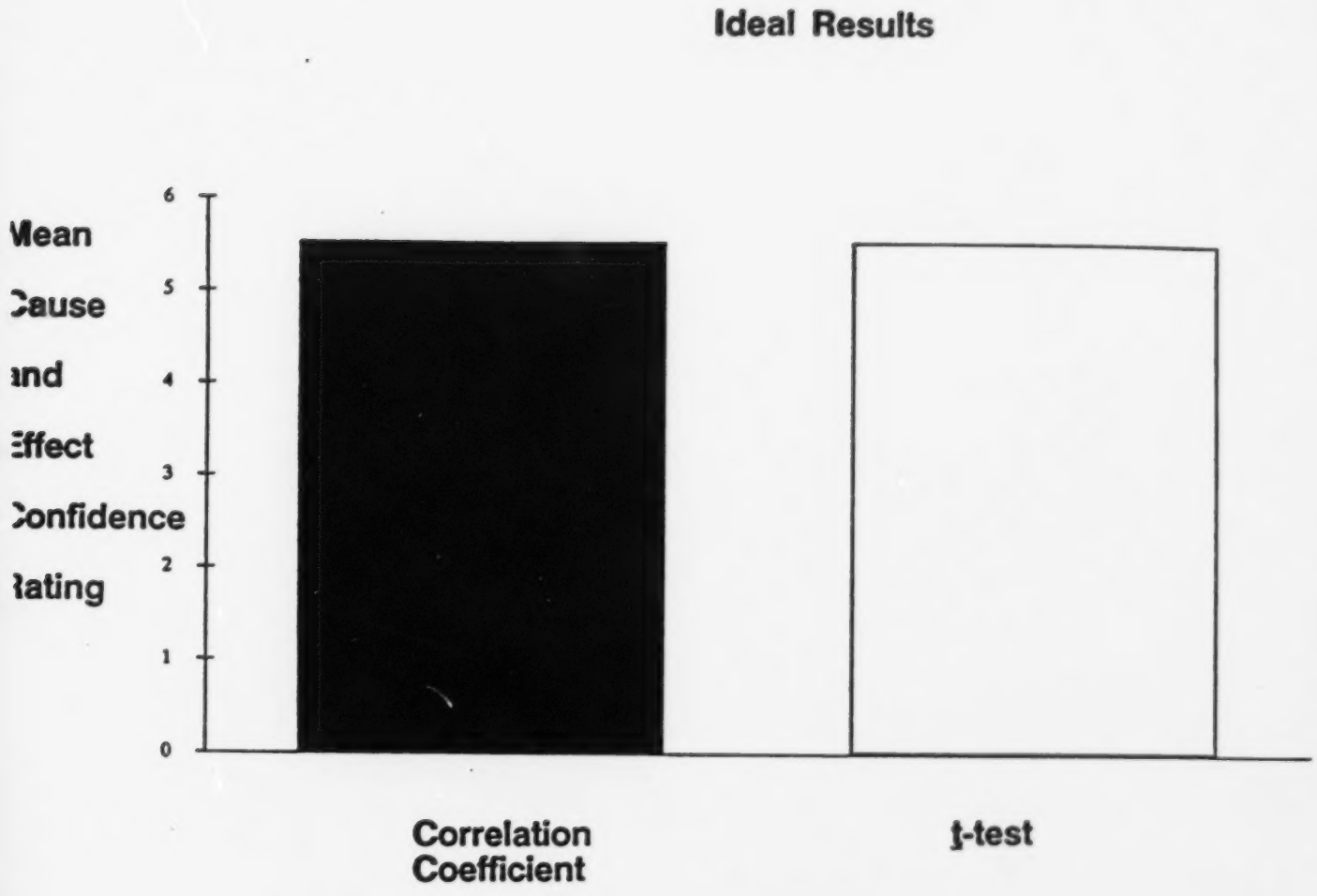

Figure 11. **Ideal Mean Ratings in Cause-And-Effect Confidence If Significance Reported with Correlation Coefficient Vs. t-test**

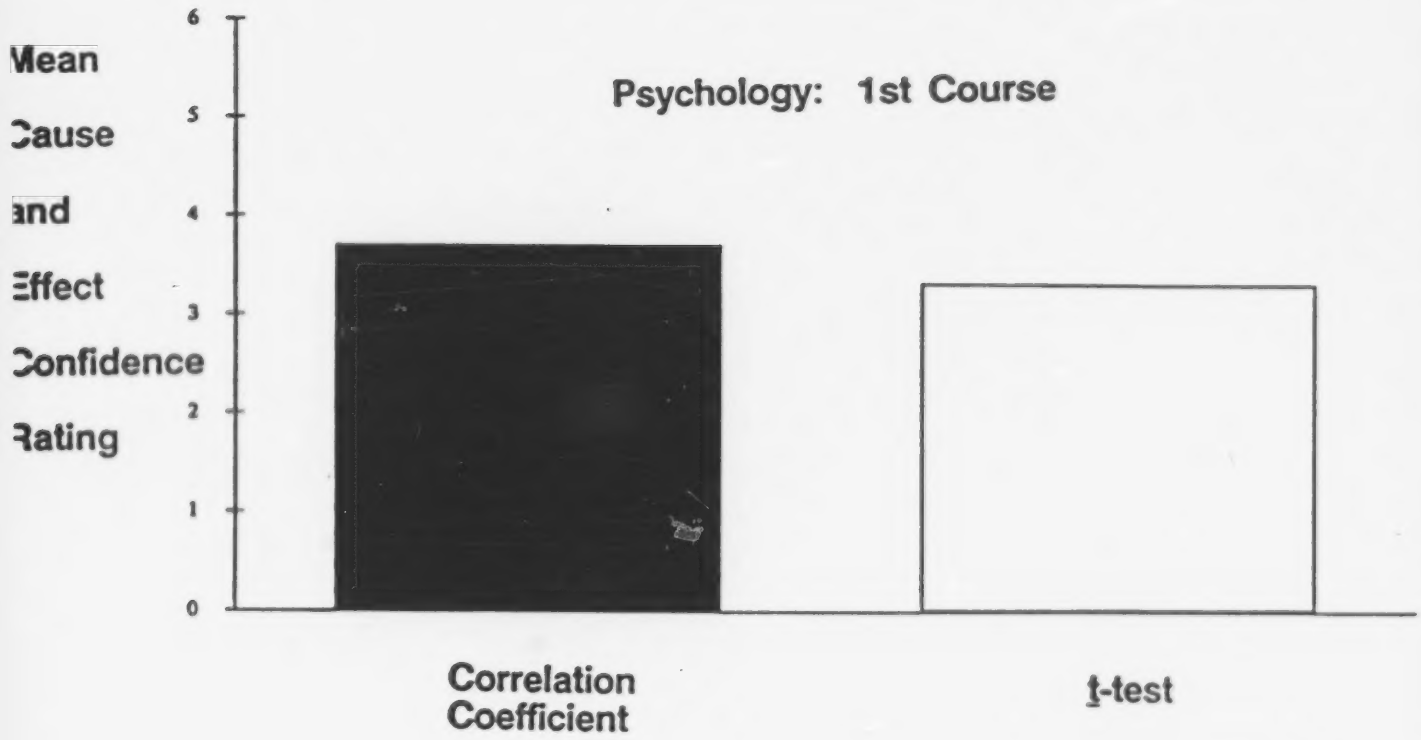

Figure 12. **Mean Cause and Effect Confidence Ratings If Significance Reported with a Correlation Coefficient Vs. a t-test: Psychology Level 1**

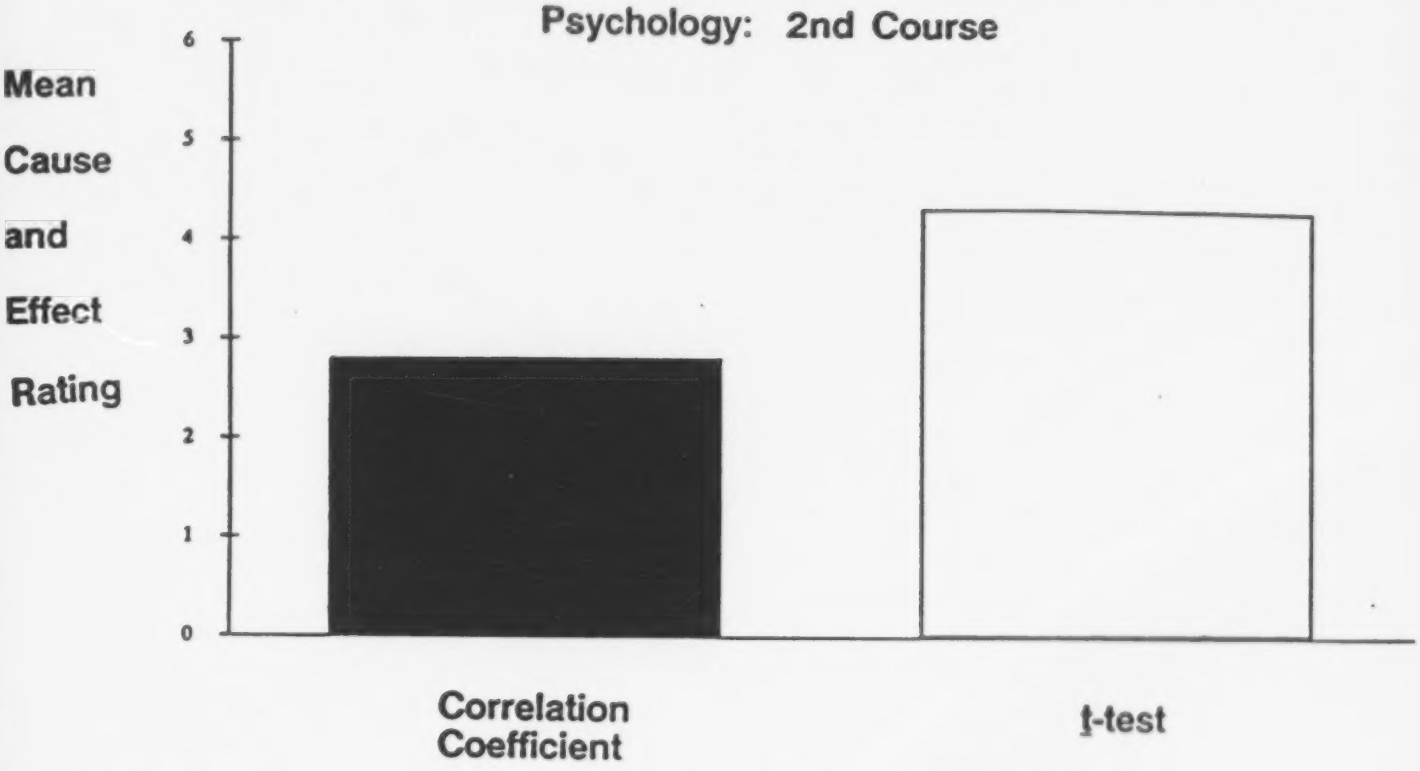

Figure 13. Mean Cause and Effect Confidence Ratings if Significance Reported with a Correlation Coefficient Vs. a t-test: Psychology Level 2

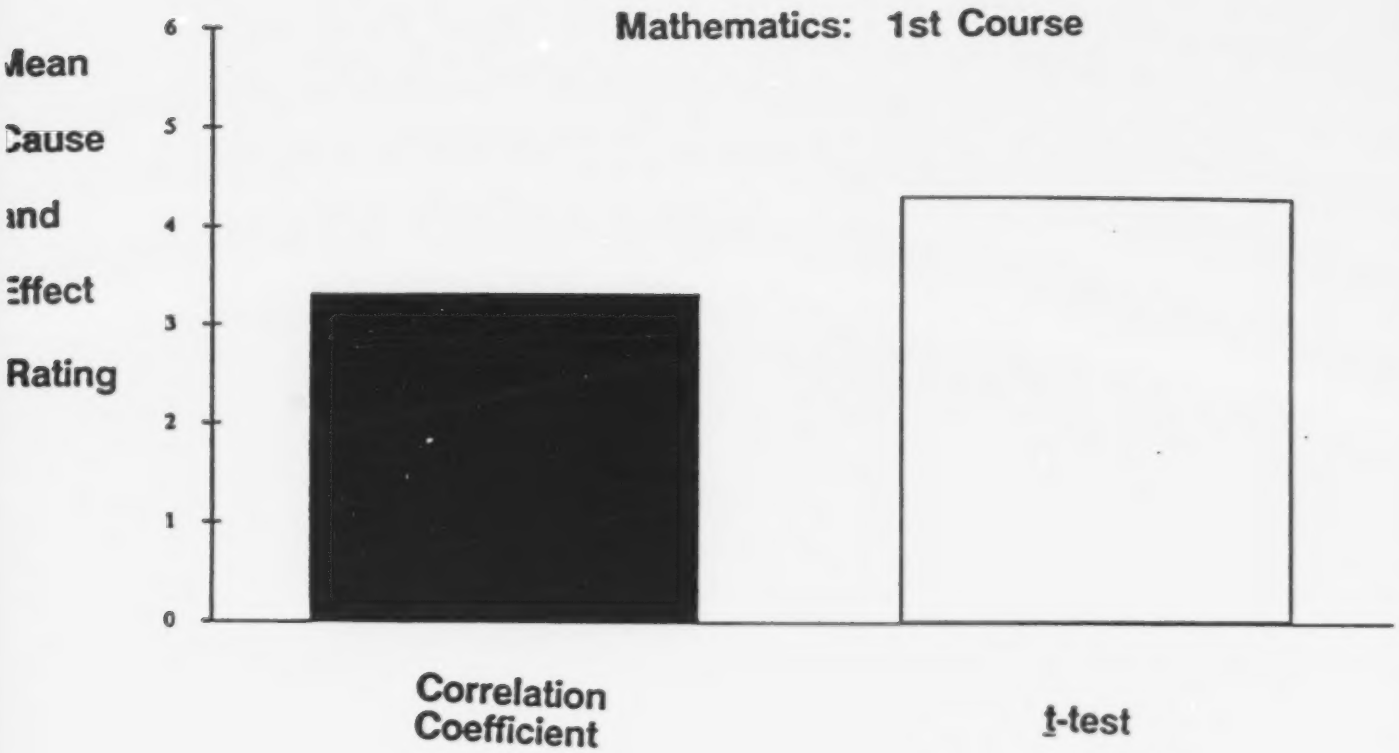

Figure 14. Mean Cause and Effect Confidence Ratings if Significance Reported with a Correlation Coefficient Vs. a t-test: Mathematics Level 1

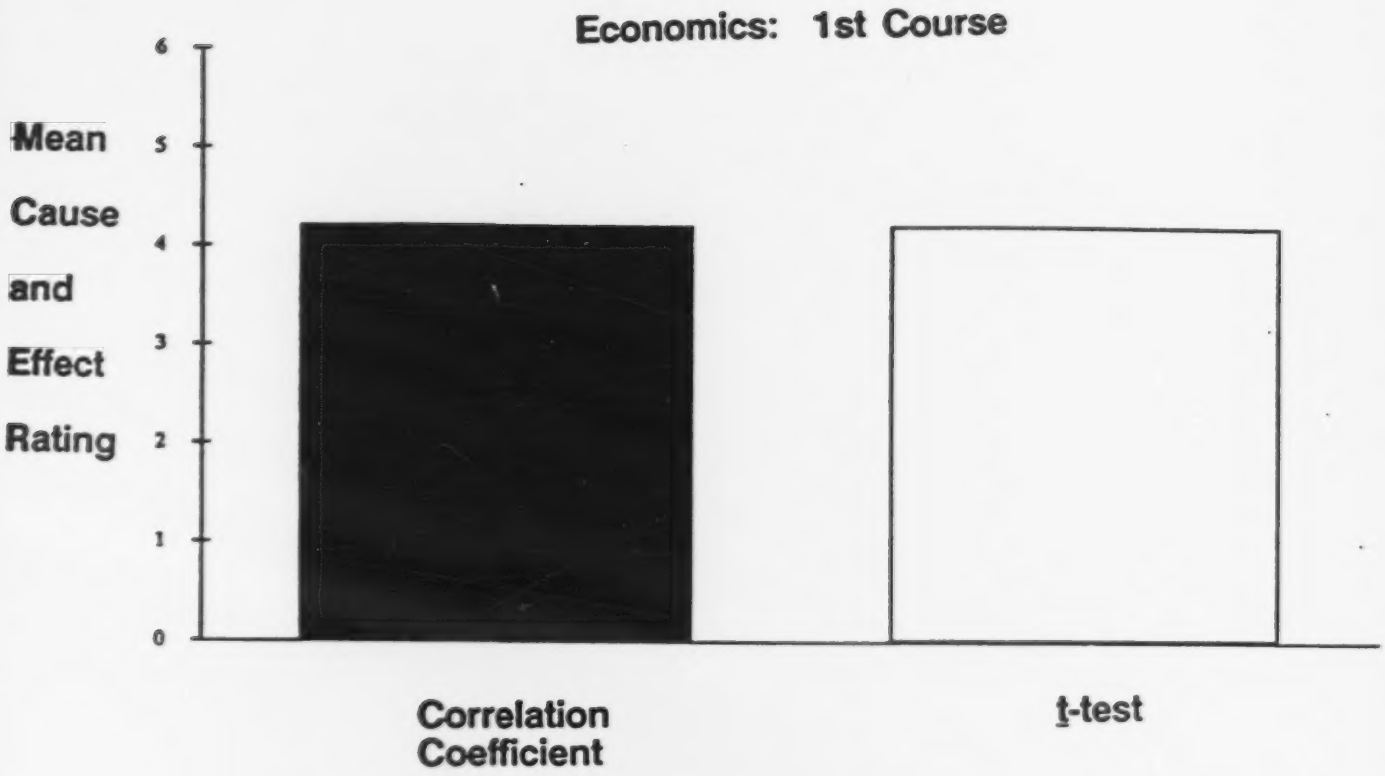

Figure 15. Mean Cause and Effect Confidence Ratings if Significance Reported with a Correlation Coefficient Vs. a t-test: Economics Level 1

Discussion

The first set of data analyses indicates that students across disciplines and universities overwhelmingly base their confidence in generalizing the results of a study on their personal beliefs about there being an independent-dependent variable relationship. This consistent reliance on the availability heuristic rather than whether the study involves a random sample or an available group is especially remarkable in light of the generalizability question being phrased as "Based on the research methodology, how confident are you in generalizing the result to other _____?" The survey booklets open so that the one-page research description is on the page directly opposite the questions. This allows students to refer back to the research description to re-examine the research method while answering the rating questions.

If students understand the importance of random sampling as a determinant of external validity, the random sampling vs. available group differences should account for the largest portion of variation in generalizability ratings. That outcome exists in only one of the 14 classes. The attention to random sampling should lead to students giving low generalizability ratings to the four studies involving nonrandom samples and high generalizability ratings to the four studies involving random samples. The mean nonrandom sampling generalizability rating for the second-level psychology class at university 3 is indeed more than the 1.2 points lower than the other 13 classes. However, the mean generalizability ratings by this class on the random sampling studies is also lower than 12 of the other 13 classes. It appears that the sensitivity to random sampling and generalizability in the one psychology class may lower the overall confidence in generalizing results even when a random sampling methodology is used.

Our hypothesis is that most students do not base their generalizability ratings on the presence of random sampling because most statistics courses across disciplines use only the random sampling model. We recommend greater emphasis on interpreting research in terms of random sampling vs. nonrandom sampling as a

means of overcoming this type of misinterpretation of research. The introduction of permutation statistics to analyze nonrandom sampling studies may also help reduce this interpretation error.

The second set of data analyses suggests that students across disciplines and universities rely heavily on the availability heuristic rather than the presence or absence of random assignment as a determinant of their confidence in drawing cause and effect conclusions. In 13 of the 14 classes, students' beliefs about the independent-dependent variable relationship account for the most variability in cause and effect ratings. This dominance of the belief variable and lack of variation accountable with the random assignment factor indicates that students are not sensitive to the importance of random assignment in drawing cause and effect conclusions.

Again, this finding is even more impressive in light of the question directing students to base their rating on the research methodology and their having the research description available as they rate cause-and-effect confidence.

If students understand the importance of random assignment as a determinant of internal validity, the random assignment vs. classificatory study differences should account for the largest percent of variation in cause and effect ratings. That outcome exists in only one of the 14 classes. The attention to random assignment should result in low cause and effect ratings for the four classificatory studies and high confidence in cause and effect for the four experimental studies. The mean cause and effect rating on classificatory studies for the second-level psychology class at university 3 is indeed 1.2 points lower than the average of the other 13 classes. However, the mean cause and effect ratings by this class on the experimental studies is also lower than each of the other 13 classes. Thus, although this one class appears to lower their confidence in cause and effect when random assignment is absent, they do not increase their confidence in cause and effect relative to other students when the studies do involve random assignment.

Despite the fact that course materials in most of these classes instruct students about the random assignment-internal validity relationship, students' beliefs reliably dominate over a reliance on research methods in making cause and effect conclusions. We recommend that courses place a heavier emphasis on interpreting research in terms of random assignment vs. classificatory groups as a means for overcoming this type of misinterpretation of research.

The outcomes of the first two sets of data analyses might lead to the discouraging conclusion that what students learn in statistics and research methods courses has little or no influence in how they interpret descriptions of research. However, the third set of data analyses indicate that students do use some of their understandings from statistics and research methods courses in their interpretation of research. Specifically, students in the second-level psychological statistics and research methods courses and students taking statistics and research methods courses in mathematics gave lower cause and effect ratings when random assignment studies were analyzed with a correlation coefficient than with a *t*-test. Unfortunately, this is an incorrect belief that the studies analyzed with a correlation statistic should not be interpreted as reflecting a cause and effect relationship. This misunderstanding probably exists in part because nonrandom-assignment, classificatory studies are referred to as correlational studies in most psychology and mathematics statistics and research methods books.

The confusion of a "correlational" research method with the correlation statistic appears to be created by the way most psychology and mathematics statistics texts deal with the correlation and causality relationship. A review of those books indicates that most do include a warning that correlating should not be interpreted as meaning causality. Unfortunately, this warning usually appears in the chapter on the correlation statistic. Thus, many students conclude that it is the correlation statistic that makes cause and effect conclusions inappropriate rather than the correct

understanding that it is the correlational or classificatory research method that should lead to reduced confidence in a cause and effect conclusion.

Economics statistics and research methods books do not place this emphasis on being cautious about drawing cause and effect conclusions with correlational studies or the correlation statistic. As indicated in the results, economics statistics and research methods students do not decrease their confidence in cause and effect as a result of data being correlation coefficient rather than a *t*-test.

Neither the first-level psychology statistics and research methods students at university 1 or 2 showed greater caution in drawing a cause and effect conclusion with the correlation statistic than with a *t*-test. Indeed, those first-level students at university 1 gave lower cause and effect ratings when data were analyzed with a *t*-test. An examination of course syllabi indicates that the caution about cause and effect interpretations with the correlation statistic is emphasized in second-level rather than first-level psychology statistics and research methods courses. The opposite-direction findings in the first-level psychology course at university 1 may be due to more familiarity with r than with t. At the time that they completed the survey, they had been tested on the correlation statistic, but not on the *t*-test.

The frequency and strength of the interpretation-of-research errors in all three sets of analyses in this study suggest the need for changes in the teaching of statistics and research methods courses. One recommendation is to acknowledge that most behavioral science research uses an available group rather than a random sampling research method. Students should learn the implications of those differences in research method both for the type of statistical analysis and for the generalizability of results. Recently published books (e.g., May, Masson, and Hunter, 1990, and Forsyth and Bohling, 1991) place a heavier emphasis on the permutation statistical model and the interpretation of available group studies.

The use of the availability heuristic rather than random assignment as a determinant of confidence in cause and effect suggests the need for students to have more experience interpreting research studies with a guide about the consequence of using random assignment to conditions vs. studying already classified subjects.

We recommend that a research methodology taxonomy such as that presented in Figure 16 be used along with interpretation-of-research assignments. This taxonomy is for studies in which the dependent variable is on an interval or ratio scale. Students would use the taxonomy to classify research studies in terms of research method and interpret the studies based on the guidelines for studies in each cell. The taxonomy should also help correct the misunderstanding about the correlation and causality relationship. Follow-up studies are needed to examine the effects of such new teaching strategies designed to overcome these interpretation-of-research errors.

CASE SELECTION

Random Sampling
INDEPENDENT VARIABLE

ASSIGNMENT	Nominal	Continuum
Random	#1 1. Significance: Random Sampling Model 2. Parameter Estimation 3. Generalize to Population by Inference 4. Cause and Effect Interpretation 5. No I.V./D.V. Functions Possible	#2 1. Significance: Random Sampling Model 2. Parameter Estimation 3. Generalize to Population by Inference 4. Cause and Effect Interpretation 5. I.V./D.V. Functions Possible
Classify	#3 1. Significance: Random Sampling Model 2. Parameter Estimation 3. Generalize to Population by Inference 4. No Cause and Effect Interpretation 5. No I.V./D.V/ Functions Possible	#4 1. Significance: Random Sampling Model 2. Parameter Estimation 3. Generalize to Population by Inference 4. No Cause and Effect Interpretation 5. I.V./D.V. Functions Possible

Available Group
INDEPENDENT VARIABLE

ASSIGNMENT	Nominal	Continuum
Random	#5 1. Significance: Permutation Model 2. No Parameter Estimation 3. Generalizations by other than Statistical Inference 4. Cause and Effect Interpretation 5. No I.V./D.V. Functions Possible	#6 1. Significance: Permutation Model 2. No Parameter Estimation 3. Generalizations by other than Statistical Inference 4. Cause and Effect Interpretation 5. I.V./D.V. Functions Possible
Classify	#7 1. Significance Permutation Model 2. No Parameter Estimation 3. Generalizations by other than Statistical Inference 4. No Cause and Effect Interpretation 5. No I.V./D.V. Functions Possible	#8 1. Significance Permutation Model 2. No Parameter Estimation 3. Generalization by other than Statistical Inference 4. No Cause and Effect Interpretation 5. I.V./D.V. Functions Possible

Figure 16. Basic Research Methodology Taxonomy Implications

References

Bohling, P.H., and Forsyth, G.A. (1990). An interpretation-of-research approach to integrating teaching, community service, and research. Paper presented at the Pennsylvania State System of Higher Education Applied Research Conference, West Chester, PA.

Bohling, P.H., and Forsyth, G.A., and May, R.B. (1990). Research methodology taxonomy and interpreting research studies. In V.P. Makosky, C.C. Sileo, L.G. Whittemore, C.P. Landry, and M.L. Skutley (Eds.) Activities Handbook for the Teaching of Psychology, Volume 3, Washington, D.C.: American Psychological Association.

Forsyth, G.A., and Bohling, P.H. (1991). Understanding Statistics and Behavioral Research Methods. Acton, Mass.: Copley Publishing Co.

Glenberg, A.M. (1988). Learning From Data: An Introduction to Statistical Reasoning. N.Y.: Harcourt Brace Jovanovich.

Keppel, G., and Zedeck, S. (1989). Data Analysis for Research Designs: Analysis of Variance and Multiple Regression/Correlation Approaches. N.Y.: W.H. Freeman and Company.

May, R.B., and Hunter, M.A. (1988). Interpreting students' interpretation of research. Teaching of Psychology, 15, 156-158.

May, R.B., and Hunter, M.A. (1990). Advantages of teaching permutation tests. Paper presented at the Annual Meetings of the Canadian Psychological Association, Ottawa, Ontario.

May, R.B., Masson, M.E.J., and Hunter, M.A. (1990). Application of Statistics in Behavioral Research. New York: Harper Collins.

Rosnow, R.L., and Rosenthal, R. (1989). Statistical procedures for the justification of knowledge in psychological science. American Psychologist, 44, 1276-1284.

Stevens, J.P. (1990). On the teaching of applied statistics and applied statistics textbooks. Paper presented at the annual meeting of the American Educational Research Association, Boston, Mass.